TITLE:

最大エントロピー定式を巡る作用素解析 : 相対エントロピーの方式 (量子確率論とエントロピー解析)

AUTHOR(S):

梅垣, 壽春

CITATION:

梅垣, 壽春. 最大エントロピー定式を巡る作用素解析 : 相対エントロピーの方式 (量子確率論とエントロピー解析). 数理解析研究所講究録 1998, 1066: 95-108

ISSUE DATE:

1998-10

URL:

http://hdl.handle.net/2433/62483

RIGHT:

京都大学学術情報リポジトリ KURENAI 紅 Kyoto University Research Information Repository

数理解析研究所講究録
1066 巻 1998 年 95-108

最大エントロピー定式を巡る作用素解析
— 相対エントロピーの方式 —

東京工業大学名誉教授
梅垣壽春

この論文の主たる目的は，前半において情報エントロピーを導入し，然る後に確率系（離散及び一般無限連続系）に於ける相対エントロピー導入と，その数理的構成を論じ，後半において絶対連続な確率変数対，更には量子統計作用素対に対して，それぞれの相対エントロピーを論じ，最大エントロピーを捉えることである．最大エントロピー定理は確率解析や量子統計力学の重要な成果として既知なものであるが，この両者を相対エントロピーの情報量概念の数学的展開によって，同一のFormulation上の定式化が可能であることを論述する．

§1　エントロピー導入と，その特性化

情報を理論的に論ずるには，情報の正確さとか不確からしさを数理的に把握することから始まる．その不確からしさの大きさを表すスカラー量をエントロピー（entropy）あるいは情報エントロピーといい，情報概念の関わる理論的根幹となっている．エントロピーはもともと熱力学上の概念で，それが情報の基本量と同一の定式によって表されること自体も興味を呼ぶことである．

X を確率事象系：

$$X = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix}, \quad \sum_{j=1}^{n} p_j = 1, 0 \le p_j \le 1$$

に対して X のエントロピー $S(X)$ は

$$S(X) = -\sum_{j=1}^{n} p_j \log p_j$$

によって与えられる．茲で $\log = \log_e, 0\log 0 = 0$.

n 個の事象のとる状態（確率分布） $p = \{p_1, \cdots, p_n\}$ の集合を Δ_n で表す．

$$\Delta_n \equiv \left\{ p = \{p_k\}_{k=1}^{n}; \sum_{k=1}^{n} p_k = 1, p_k \geq 0 \right\}$$

である．目標の定式化と公理系は，集合 $\Delta = \coprod \Delta_n$ 上で定義され半直線 $[0,\infty)$ に値をとる関数 $S(\cdot)$ に関して次のように与えられる：

$\langle\langle S \rangle\rangle$ （Shannon のエントロピー定式）

$$S(p) = S(p_1,\cdots,p_n) = -\sum_{k=1}^{n} p_k \log p_k .$$

$\langle\langle SK \rangle\rangle$ （Shannon - Khinchin の公理系）

$\langle SK1 \rangle$ $S(\cdot)$ は各 Δ_n 上の連続関数であり，かつその上で最大値をとる：

$$\max\{S(p); p \in \Delta_n\} = S(1/n,\cdots,1/n) > 0,$$

$\langle SK2 \rangle$ $S(p_1,\cdots,p_n,0) = S(p_1,\cdots,p_n)$,

$\langle SK3 \rangle$ $S(q_{11},\cdots,q_{1m_1},\cdots,q_{n1},\cdots,q_{nm_n})$

$$= S(p_1,\cdots,p_n) + \sum_{k=1}^{n} p_k S(q_{k1}/p_k,\cdots,q_{km_k}/p_k).$$

ここで，$p_k = \sum_{j=1}^{m_k} q_{kj} > 0 \ (m_k \geq 2, k = 1,n)$ とする．

$\langle\langle F \rangle\rangle$ （Faddeev の公理系）

$\langle F1 \rangle$ 関数 $f(p) = S(p,1-p)$ は $0 \leq p \leq 1$ 上で連続かつ少なくとも 1 点 p_0 で

$$f(p_0) > 0,$$

$\langle F2 \rangle$ $(p_1,\cdots,p_n) \in \Delta_n$ の任意の置換 $(p_1',\cdots,p_n')$ に対して不変：

$$S(p_1,\cdots,p_n) = S(p_1',\cdots,p_n'),$$

$\langle F3 \rangle$ $p_n = q + r, q > 0, r > 0$ のとき

$$S(p_1,\cdots,p_{n-1},q,r) = S(p_1,\cdots,p_n) + p_n S(q/p_n, r/p_n).$$

$\langle\langle SK \rangle\rangle$ は Shannon が発見し導入，Khinchin がそれを分かりやすくし，証明も簡単にしたもので，その後 Faddev によりさらに簡潔な $\langle\langle F \rangle\rangle$ が示された．それぞれ対応する文献（出典）は Shannon [10]（for $\langle\langle S \rangle\rangle$），Khinchin [4]（for $\langle\langle S \rangle\rangle$, $\langle\langle SK \rangle\rangle$），Faddeev [1]（for $\langle\langle F \rangle\rangle$），更に邦文 [3]，[9]，[14]，[15] を参照した．

§ 2　エントロピー導入と，その特性化

本論に於いて，基本的役割を演出する基本不等式を示そう．

1°．対数不等式

$$\log t \geq 1 - t^{-1} \quad (t > 0),$$

等号の成立 $\Leftrightarrow t=1$，茲で $\log=\log_e$ とする（以下同様）．

2°． $t\log t \geq t-1\,(t\geq 0)$,

等号の成立 $\Leftrightarrow t=1$，茲で $0\log 0=0$ とする（以下同様）．

3°．Kullback-Leibler（K-L）不等式

$$\int_\Omega f(\omega)\log g(\omega)d\mu(\omega) \leq \int_\Omega f(\omega)\log f(\omega)d\mu(\omega)$$

for $f,g\in L^1(\Omega)^+, \|f\|_1=\|g\|_1, \operatorname{supp} f\subset \operatorname{supp} g$,

等号の成立 $\Leftrightarrow f=g\,(\mu-\text{a.e.})$,

茲で，$\Omega=(\Omega,\mathfrak{L}_\Omega,\mu)$ は σ-有限測度空間，$\operatorname{supp} f$ は函数 f の台，即ち $\operatorname{supp} f=\{\omega\in\Omega; f(\omega)>0\}$.

4°．可算確率分布の対 $\mathbf{p},\mathbf{q}$

$$\mathbf{p}=(p_1,p_2,\cdots),\mathbf{q}=(q_1,q_2,\cdots)$$

即ち，$p_j,q_j\geq 0\,(j=1,2,\cdots)$, $\sum_{j=1}^\infty p_j=\sum_{j=1}^\infty q_j=1$ である $\mathbf{p},\mathbf{q}$ に対して

$$\sum p_j\log p_j \geq \sum p_j\log q_j \tag{2.1}$$

等号の成立 $\Leftrightarrow p_j=q_j\,(j=1,2,\cdots)$.

この不等式は Shannon が同時エントロピーを論じ，更に情報路（Channel）の伝送速度を導入した際に基本不等式としてcreat した不等式であり，また 3° の Kullback-Leibler 不等式と表示したものもこの意味で Shannon の基本不等式の帰結である．

1° の証明．これは初等数学の範囲である：

$y\overset{\Delta}{=}\log t-(1-t^{-1})=\log t+t^{-1}-1\,(t>0)$,

$y'\left(=\dfrac{dy}{dt}\right)=t^{-1}-t^{-2}=(t-1)t^{-2}$,

t	0	↗	1	↗
y'		−	0	+
y''		↘		↗

故に $y=0\Leftrightarrow y'=0\Leftrightarrow t=1$.

2° の証明．不等式 2° ⇔ 1° であるから 2° は自明．

3° の証明．不等式 2° を用いることによって，$\operatorname{supp} f\subset\operatorname{supp} g$ より

$$\frac{f}{g}\log\frac{f}{g} \geq \frac{f}{g} - 1 \quad (\text{on} \quad \operatorname{supp} g) \tag{2.2}$$

故に

$$f\log f \geq f\log g + f - g \quad \text{a.e. in } \boldsymbol{\Omega}.$$

この両辺を積分することによって目標の不等式を得る．更に 2° によって，不等式 (1.2) が等号 $\Leftrightarrow$ $f = g$ (a.e.).

4° は 3° の特別の場合である．

《註》3° の K-L 不等式は次の § 3 で役割を演じ，その効果は本論の全般に呈る．更に本論に於いては一部分のみに関与する程度であるが，函数解析・量子力学・確率論その他広汎に呈って常に重要な役割を演ずる Schwartz の不等式を付記しておく．

測度空間 Ω $(=(\Omega, \mathfrak{L}_\Omega, \mu)$， 3° 参照) 上の函数 Hilbert 空間 $L^2(\Omega)$ を設定する．このとき $\forall f, g \in L^2(\Omega)$ に対して Schwartz の不等式が次の 3 つの定式によって表される．

$$\left|\int_\Omega \overline{f(\omega)} g(\omega) d\mu(\omega)\right|^2 \leq \int |f(\omega)|^2 d\mu(\omega) \cdot \int |g(\omega)|^2 d\mu(\omega) \tag{2.3}$$

これを内積とノルムを用いて表すと

$$|\langle f, g\rangle| \leq \|f\|_2 \cdot \|g\|_2. \tag{2.4}$$

更に $\Omega = (\Omega, \mathfrak{L}_\Omega, \mu)$ が確率空間の場合は期待値の記号 $E(\cdot)$ を用いて

$$\left|E(\bar{f} g)\right|^2 \leq \left|E(|f|^2) \cdot E(|g|^2)\right| \tag{2.5}$$

と表す．本節に於ける主たる参考文献は [3,5 ～ 9,13,14] である．

§ 3　Kullback-Leibler (K-L) 情報量

Shannon [10] の情報エントロピー発見（1948年）の 5 年後，Kullback と Leibler [5] が新たな情報量を導入した．これを，先ず可算事象系の場合に設定しよう：

$$\mathcal{A} = (a_1, a_2, \cdots),$$

$$\Delta_\infty \stackrel{\Delta}{=} \left\{ \mathbf{p} = (p_1, p_2, \cdots); \sum_{j=1}^{\infty} p_j = 1, p_j \geq 0 \right\}$$

とする．対 $\mathbf{p}, \mathbf{q} \in \Delta_\infty$ に対して完全事象系の対

$$\begin{pmatrix} \mathcal{A} \\ \mathbf{p} \end{pmatrix}, \begin{pmatrix} \mathcal{A} \\ \mathbf{q} \end{pmatrix}$$

が与えられるとき，$\begin{pmatrix} \mathcal{A} \\ \mathbf{p} \end{pmatrix}, \begin{pmatrix} \mathcal{A} \\ \mathbf{q} \end{pmatrix}$（or $\mathbf{p}, \mathbf{q}$）の K-L 情報量が

$$S(\mathbf{p}/\mathbf{q}) = \sum_{j=1}^{\infty} \left(p_j \log p_j - p_j \log q_j \right) \tag{3.1}$$

によって定義される．最近，これは相対エントロピーの名称で呼ばれている．

これに関して，次の定理が得られる．

定理 3.1　可算確率分布の対 $\mathbf{p}, \mathbf{q}$ に対し

$$S(\mathbf{p}/\mathbf{q}) \geq 0\ ; 等号 \Leftrightarrow \mathbf{p} = \mathbf{q}$$

ここで $\mathbf{p} = \mathbf{q}$ は $p_j = q_j\ (\forall j)$ の意味．

この定理は不等式 (2.1) によって自明．また，$\mathbf{p}, \mathbf{q}$ は可算，つまり $\mathbf{p}, \mathbf{q} \in \Delta_\infty$ の場合であるが，勿論 $\mathbf{p}, \mathbf{q} \in \Delta_n$ の場合も同様である．

《註》この New Concept は Kullback と Leibler [5] が初めに目標としていた数理統計学の分野が Operations Research に適用，発展し，さらに経営情報などの分野にも及び，集積された製品の山から不良品の選別をする際などで最大情報量の計算を用いて活動される．最大情報量は本論の後半で相対エントロピー解析として厳密に展開する．

§ 4．確率空間上の相対エントロピー

これまでは離散的な確率事象について論じて来たが，暫くの間は必ずしも離散的でない，一般の場合について論ずる．

$\Omega=(\Omega,\mathcal{L},\mu)$ を一般の確率測度空間とし，上述の $K-L$ 情報量 $S(\mathbf{p}/\mathbf{q})$ が Ω 上に次のように Ω 上に展開される：

$$AC(\Omega)\equiv\{\xi;\ \xi \text{ は } \Omega \text{ 上の実数値確率変数であり，且つ絶対連続である }\} \tag{4.1}$$

全ての対 $\xi,\eta\in AC(\Omega)$ に対して密度函数（Radon-Nikodym 微分）が次のように与えられる：

$$p_\xi(t)\equiv\frac{\mu\left(\xi^{-1}(dt)\right)}{dt},\ p_\eta(t)\equiv\frac{\mu\left(\eta^{-1}(dt)\right)}{dt}.$$

このとき，離散的な場合（§ 3，(3.1)）が次の積分計算によって拡張される：

$$S(\xi/\eta)=\int_{-\infty}^{\infty}p_\xi(t)\left(\log p_\xi(t)-\log p_\eta(t)\right)dt. \tag{4.2}$$

これを対 ξ,η の<u>相対エントロピー</u>という．このとき § 2 の $K-L$ 不等式 3° により，直ちに次の定理を得る：

定理 4.1　任意の対 $\xi,\eta\in AC(\Omega)$ に対して

$$S(\xi/\eta)\geq 0;\quad =0\ \Leftrightarrow\ p_\xi=p_\eta.$$

《註》pair でなく single $\xi\in AC(\Omega)$ に対しては

$$S(\xi)=S\left(p_\xi\right)=-\int_{-\infty}^{\infty}p_\xi(t)\log p_\xi(t)dt \tag{4.3}$$

で表す．これは Shannon エントロピーである．

ここで本 § 4 の主定理である最大エントロピーの特性化に向かって進めよう．
全ての $\xi\in AC(\Omega)$ に対して基本スカラー量

$$\xi \text{ の期待値 } =E(\xi)$$
$$\xi \text{ の分散 } =\mathrm{var}(\xi)=E\left(\xi^2\right)-E(\xi)^2$$

などが定義される．このとき定数の対 $m,\sigma>0$ に対して

$$\mathcal{G}_{m,\sigma}\equiv\left\{\xi\in AC(\Omega);E(\xi)=m,\mathrm{var}(\xi)=\sigma^2\right\}.$$

この対 m,σ に対してGaussian 確率変数 $G\left(\text{with } E(G)=m, \mathrm{var}(G)=\sigma^2\right)$ は $G\in\mathcal{G}_{m,\sigma}$ である．この G は

$$p_G(t)=\left(2\pi\sigma^2\right)^{-1/2}\exp\left(-(t-m)^2/2\sigma^2\right)$$

であり，従って G のエントロピーは

$$\begin{aligned}S(G)&=-\int_{-\infty}^{\infty}p_G(t)\log\left((2\pi\sigma)^{-1/2}\exp\left(-(t-m)^2/2\sigma^2\right)\right)dt\\&=\frac{1}{2}\log\left(2\pi e\sigma^2\right).\end{aligned}\tag{4.4}$$

以上の準備の下で次の定理が成り立つ．

定理 4.2（最大エントロピー定理 1）
任意の $\xi\in\mathcal{G}_{m,\sigma}$ に対して，次の (i) ～ (iv) が成立：
(i) $S(G)=-\int_{-\infty}^{\infty}p_\xi(t)\log p_G(t)dt$
(ii) $S(G)-S(\xi)=S(\xi/G)$
(iii) $S(\xi/G)=0\Leftrightarrow\xi=G$
(iv) 与えられた一つの $\xi_0\in\mathcal{G}_{m,\sigma}$ に対して

$$\begin{aligned}&S(\xi_0)=\max\left\{S(\xi);\xi\in\mathcal{G}_{m,\sigma}\right\}\\&\Leftrightarrow\quad \xi_0=G\end{aligned}$$

証明 (i) $\begin{aligned}E\left(\xi^2\right)&=\int|\xi(\omega)|^2d\mu(\omega)\\&=\int t^2\mu\left(\xi^{-1}(dt)\right)=\int t^2p(t)dt.\end{aligned}$

一方，$E(\xi)=\int t\cdot p_\xi(t)dt$ が成立するから

$$\sigma^2=\int t^2p_\xi(t)dt-\left(\int t\cdot p_\xi(t)dt\right)^2,$$

更に

$$\int p_G(t)(t-m)^2dt=\int p_\xi(t)(t-m)^2dt=\sigma^2.$$

一方

$$\begin{aligned}\sigma^2-\int p_\xi(t)\log p_G(t)dt&=\int p_\xi(t)\left[\frac{1}{2}\log\left(2\pi\sigma^2\right)+\frac{(t-m)^2}{2\sigma^2}\right]dt\\&=\frac{1}{2}\log\left(2\pi\sigma^2\right)+\frac{\sigma^2}{2\sigma^2}\\&=\frac{1}{2}\log\left(2\pi e\sigma^2\right)=S(G).\end{aligned}$$

従って (i) が成立．

(ii) ： (i) によって

$$\begin{aligned} S(G)-S(\xi) &= \int p_\xi(t)\log p_\xi(t)dt - \int p_G(t)\log p_G(t)dt \\ &= \int p_\xi(t)\log p_\xi(t)dt - \int p_\xi(t)\log p_G(t)dt \\ &= S(\xi/G) \end{aligned}$$

(iii) は定理４.１によって，(iv) は (ii)，(iii) を用いて直ぐ云える.

§５．GKY－相対エントロピー

前節迄は絶対連続な確率変数の対の相対エントロピーなどを論じたが，ここでは確率測度の対を与え，その測度間の絶対連続性の有無に関する判定条件として相対エントロピーの導入とその特性化について論ずる．表記の GKY は発見者 Gelfand，Kolmogorov，それに Yaglom の頭文字を表している．これは数学的に見事な完成された結果で，確率変数の対や，同測度対についての相対エントロピーを論ずる際には，常に登場せねばならぬ重要な定理である．

与えられた可測空間 $(\Omega,\mathcal{L}_\Omega)$ に於いて，
Ω の有限部分集合族 $\mathcal{F}=\{F_1,F_2,\cdots,F_n\}$ が $\mathcal{L}_\Omega$－可測分割であるとは，

$$F_i\cap F_j\neq\varnothing\ (i\neq j),\ \bigcup_{j=1}^{n}F_j=\Omega,,\ \mathcal{F}\subset\mathcal{L}_\Omega$$

のときをいい，これによって

$$\mathcal{P}(\Omega,\mathcal{L}_\Omega)\equiv\{\mathcal{F};\mathcal{F}\text{ は }\Omega\text{ の有限 かつ }\mathcal{L}_\Omega-\text{可測分割 }\}$$

と置く.

このとき，確率測度の任意の対 μ,ν に対して

$$S(\mu/\nu)\equiv\sup\left\{\sum_{F\in\mathcal{F}}\mu(F)\log\frac{\mu(F)}{\nu(F)};F\in\mathcal{P}(\Omega,\mathcal{L}_\Omega)\right\}$$

を GKY－相対エントロピーという．このとき，次の様に，測度論上の有効な定理を得る.

定理５.１（GKY 定理）　可測空間 $(\Omega,\mathcal{L}_\Omega)$ 上の任意の確率測度の対 μ,ν に対して，次の条件 (i) と (ii) が成立する：

(i)　$\mu<<\nu$（μ は ν に関して絶対連続）のとき

$$S(\mu/\nu)=\int_\Omega\left(\frac{d\mu}{d\nu}\log\frac{d\mu}{d\nu}\right)d\nu=\int_\Omega\left(\log\frac{d\mu}{d\nu}\right)d\nu$$

(ii) $\mu \not\ll \nu$ のとき $S(\mu/\nu)=+\infty$.

これの証明は Kallianpour [2] の方法が知られており，邦文では文献 [14]，[15] を参照され度い．この定理の効用は函数解析・確率論に於いて顕著である．

§ 6．作用素対の相対エントロピー

この § を通じて作用素対 A, B を次の様に設定する． A, B を Quantum - Statistical Operators，i.e.，A, B は可分 Hilbelt 空間 $\mathcal{H}$ 上の正値線形作用素（i.e.，$A,B\geq 0$ ）で完全連続且つ trace 1 (: $\mathrm{Tr}(A)=\mathrm{Tr}(B)=1$) とする [7]，[8]，[11]．このとき A, B は von Neumann - Schatten 素子（ [11] ）によって

$$A=\sum_{j=1}^{\infty}\alpha_j x_j\otimes\bar{x}_j,\ B=\sum_{k=1}^{\infty}\beta_k y_k\otimes\bar{y}_k \tag{6.1}$$

茲で

$$Ax_j=\alpha_j x_j \ \text{ and } \ By_k=\beta_k y_k \quad (\forall j,k)$$

$$\sum\alpha_j=\sum\beta_k=1,\quad \alpha_j\downarrow 0,\beta_k\downarrow 0$$

（ α_j は A の固有値，x_j は固有ベクトル，対 β_k, y_k も同様）．以下説明を単略化するため $\{x_j\},\{y_k\}$ は共に CONS in $\mathcal{H}$ とする．

定義 6.1　次の様に２種類のエントロピーを定義する：
作用素 A のエントロピー：$S(A)=-\mathrm{Tr}(A\log A)$
A の B に関する（or 対 A,B の）相対エントロピー：

$$S(A/B)=\mathrm{Tr}(A\log A-A\log B)$$

Remark.　$S(A)$ は von Neumann [8] によって introduce され，$S(A/B)$ は $K-L$ 情報量の非可換展開として Umegaki [12]，[13] によって，より一般の von Neumann Algebras 上に導入された．

定理 6.1　両エントロピーは次の様に級数展開表示される．
(1°)　$S(A)=-\sum_j\alpha_j\log\alpha_j$

(2°) $S(A/B)=\sum_{j,k}\alpha_j|\langle x_j,y_k\rangle|^2(\log\alpha_j-\log\beta_k)$

ここで $\sum_j=\sum_{j=1}^{\infty}$, $\sum_{j,k}=\sum_{j=1}^{\infty}\sum_{k=1}^{\infty}$

証明 (1°) (6.1) によって

$$\begin{aligned}A\log A&=\left(\sum_j\alpha_j x_j\otimes\bar{x}_j\right)\left(\sum_k(\log\alpha_k)x_k\otimes\bar{x}_k\right)\\&=\sum_{j,k}\alpha_j\log\alpha_k(x_j\otimes\bar{x}_j)(x_k\otimes\bar{x}_k)\\&=\sum_{j,k}(\alpha_j\log\alpha_k)\langle x_j,x_k\rangle x_j\otimes\bar{x}_k\\&=\sum_j(\alpha_j\log\alpha_j)x_j\otimes\bar{x}_j\end{aligned}\tag{6.2}$$

$$\therefore S(A)=-\mathrm{Tr}(A\log A)=-\sum_j\alpha_j\log\alpha_j$$

(2°) 再び (6.1) によって

$$\begin{aligned}A\log B&=\left(\sum_j\alpha_j x_j\otimes\bar{x}_j\right)\left(\sum_k(\log\beta_k)y_k\otimes\bar{y}_k\right)\\&=\sum_{j,k}(\alpha_j\log\beta_k)(x_j\otimes\bar{x}_j)(y_k\otimes\bar{y}_k)\\&=\sum_{j,k}(\alpha_j\log\beta_k)\langle x_j,y_k\rangle x_j\otimes\bar{y}_k\end{aligned}$$

$$\begin{aligned}\mathrm{Tr}(A\log B)&=\sum_{j,k}(\alpha_j\log\beta_k)\langle x_j,y_k\rangle\mathrm{Tr}(x_j\otimes\bar{y}_k)\\&=\sum_{j,k}\alpha_j|\langle x_j,y_k\rangle|^2\log\beta_k\end{aligned}\tag{6.3}$$

$$\begin{aligned}\therefore S(A/B)&=\mathrm{Tr}(A\log A-A\log B)\\&=\sum_{j,k}\alpha_j|\langle x_j,y_k\rangle|^2(\log\alpha_j-\log\beta_k)\end{aligned}\tag{6.4}$$

(Q.E.D.)

定理 6.2 作用素の対 A,B に対して次の (1°) ~ (4°) が成立する：

(1°) $S(A)\geq 0$,

(2°) $S(A)=0\Leftrightarrow A=x_{j_0}\otimes\bar{x}_{j_0}$ for some j_0,

(3°) $S(A/B)\geq 0$,

(4°) $S(A/B)=0\Leftrightarrow A=B$.

証明 (1°) is clear ; (2°) $S(A)=0$

$$\Leftrightarrow\exists j_0:\alpha_{j_0}=1,\alpha_j=0\ (\forall j\neq j_0)$$
$$\Leftrightarrow A=x_{j_0}\otimes\bar{x}_{j_0};$$

(3°) 等式 (6.3) , (6.4) によって

$$S(A/B) = \sum_j \alpha_j \log \alpha_j - \sum_j \alpha_j \left(\sum_k \left| \langle x_j, y_k \rangle \right|^2 \log \beta_k \right)$$

$$\geq \sum_j \alpha_j \log \alpha_j - \sum_j \alpha_j \left(\log \sum_k \beta_k \left| \langle x_j, y_k \rangle \right|^2 \right)$$

（$\sum_j \left| \langle x_j, y_k \rangle \right|^2 = 1$ と log の凹性を用いた）

$$= \sum_j \alpha_j \log \alpha_j - \sum \alpha_j \log \langle x_j, Bx_j \rangle$$

$$= \sum \alpha_j \log \left(\alpha_j / \langle x_j, Bx_j \rangle \right)$$

$$= \sum \langle x_j, Bx_j \rangle \left(\frac{\alpha_j}{\langle x_j, Bx_j \rangle} \log \frac{\alpha_j}{\langle x_j, Bx_j \rangle} \right)$$

不等式 $t \log t \geq t - 1 \, (t > 0)$ を用いて

$$\geq \sum \langle x_j, Bx_j \rangle \left(\frac{\alpha_j}{\langle x_j, Bx_j \rangle} - 1 \right)$$

$$= \sum \alpha_j - \sum \langle x_j, Bx_j \rangle = 1 - 1 = 0 \tag{6.5}$$

∴（3^0）を得る.

（4°）(6.5) に至る不等式の series に於いて

$$\sum_j \alpha_j \left(\log \sum \beta_k \left| \langle x_j, y_k \rangle \right|^2 \right) \geq \sum_j \alpha_j \left(\sum_k \left| \langle x_j, y_k \rangle \right|^2 \log \beta_k \right) \tag{6.6}$$

∴ $S(A/B) = 0 \Rightarrow$ 不等式 (6.6) は等式となる

$$\therefore \log \left(\sum_k \beta_k \left| \langle x_j, y_k \rangle \right|^2 \right) = \sum_k \left| \langle x_j, y_k \rangle \right|^2 \log \beta_k \quad (\forall j)$$

$$\therefore \forall j \, \exists k_j \text{ and const } \gamma_j ; x_j = \gamma_j y_{k_j}, \left| \gamma_j \right| = 1$$

$$\therefore A = B.$$

《註》作用素対 A, B の相対エントロピー $S(A/B)$ を論じて来たが，当然所論から明らかなように無限級数和は $S(A/B) = +\infty$ である場合も含めて well-defined であることも主張に含まれている.

§４で論じた確率変数に於ける最大エントロピー定理１は，ここで対象にしてい

る物理系に於いても全く同様な定式に展開される．対象としている物理系の Hamiltonian を H とする．これは自己共役作用素である．これと逆温度

$$\beta = \frac{1}{kT} \quad \begin{pmatrix} k = \text{Boltzmann Const} \\ T = \text{絶対温度} \end{pmatrix}$$

に対して自己共役作用素 $e^{-\beta H}$ が有限 trace をもつとき密度作用素

$$\mathcal{G} = \frac{1}{\mathrm{Tr}\left(e^{-\beta H}\right)} e^{-\beta H}$$

が Gibbs state といわれ，量子系に於ける平衡状態を表している．このとき§４の最大エントロピー定式が Gaussian G によって捉えられたと全く同様な一連の定式が得られる．

$$\mathbf{Q}_H \equiv \left\{B; \mathrm{Tr}(BH) = \mathrm{Tr}(\mathcal{G}\, H) \text{ を満たす量子統計作用素}\right\}$$

とするとき目標の定理：

定理 6.3（最大エントロピー定理２） 任意の作用素 $B \in \mathbf{Q}_H$ に対して，次の (i) ～ (iv) が成立する：

(i) $S(\mathcal{G}) = -\mathrm{Tr}(B \log \mathcal{G})$,

(ii) $S(\mathcal{G}) - S(B) = S(B/\mathcal{G}) \geq 0$,

(iii) $S(\mathcal{G}/B) = 0 \Leftrightarrow S(B/\mathcal{G}) = 0 \Leftrightarrow B = \mathcal{G}$,

(iv) 一つの作用素 $B_0 \in \mathbf{Q}_H$ に対して

$$S(B_0) = \max\left\{S(B); B \in \mathbf{Q}_H\right\} \Leftrightarrow B_0 = \mathcal{G}_0$$

証明 (i)：
$$\begin{aligned}
\mathrm{Tr}(\mathcal{G} \log \mathcal{G}) &= \mathrm{Tr}\left(e^{-\beta H} / \mathrm{Tr}\left(e^{-\beta H}\right)\right) \log\left(e^{-\beta H} / \mathrm{Tr}\left(e^{-\beta H}\right)\right) \\
&= \mathrm{Tr}\left(\left(e^{-\beta H} / \mathrm{Tr}\left(e^{-\beta H}\right)\right)\left(-\beta H - \log\left(\mathrm{Tr}\left(e^{-\beta H}\right)\right)\right)\right) \\
&= -\mathrm{Tr}\left(e^{-\beta H}\right)^{-1} \mathrm{Tr}\left(e^{-\beta H}\left(\beta H + \log \mathrm{Tr}\left(e^{-\beta H}\right)\right)\right) \\
&= -\mathrm{Tr}\left(e^{-\beta H}\right)^{-1}\left(\beta \mathrm{Tr}\left(e^{-\beta H} \cdot H\right) + \mathrm{Tr}\left(e^{-\beta H} \log \mathrm{Tr}\left(e^{-\beta H}\right)\right)\right) \\
&= -\beta \mathrm{Tr}\left(e^{-\beta H}\right)^{-1} \mathrm{Tr}\left(e^{-\beta H} \cdot H\right) - \log \mathrm{Tr}\left(e^{-\beta H}\right) \\
&= -\beta \mathrm{Tr}(\mathcal{G}\, H) - \log \mathrm{Tr}\left(e^{-\beta H}\right) \\
&= -\beta \mathrm{Tr}(BH) - \log \mathrm{Tr}\left(e^{-\beta H}\right) \\
&= -\mathrm{Tr}\left(B(\beta H)\right) - \mathrm{Tr}(B) \log \mathrm{Tr}\left(e^{-\beta H}\right) \\
&= \mathrm{Tr}\left(B \log e^{-\beta H}\right) - \mathrm{Tr}(B) \log \mathrm{Tr}\left(e^{-\beta H}\right)
\end{aligned}$$

$$= \mathrm{Tr}\left(B\log\left(e^{-\beta H}/\mathrm{Tr}\left(e^{-\beta H}\right)\right)\right)$$
$$= \mathrm{Tr}\left(B\log\mathcal{G}\right)$$

即ち (i) を得る．(ii) は次のことより

$$0 \le S(B/\mathcal{G}) = \mathrm{Tr}(B\log B - B\log\mathcal{G}) = -S(B) + S(\mathcal{G}).$$

(iii) は定理６．２より，最後に (iv) は，これらの結果より自明．

Bibliography

[1] A.D.Faddeev , On the notion of entropy of a finite probability space , Uspehi , Mat Nauvs 11 (1958) , 227-236.

[2] G.Kallianpour , On the amount of information contained in a σ-field , in Contribute to Prob. & Statist , Stanford Univ. Press. (1960).

[3] 国澤清典・梅垣壽春 , 情報理論の進歩 , 岩波書店 , 1960年.

[4] L.Khinchin , Mathematical Fundation of Information Theory , Dover (1989) (English Transl. from YMH 8 (1953) and 11 (1956)).

[5] S.Kullback and P.A.Leibler , On information and sufficiency , Ann. Math. Statistics , 22 (1951) , 79-86.

[6] M.Nakamura and H.Umegaki , A note on the entropy for operator algebra , Proc. Japan Acad. 37 (1961) , 149-156.

[7] M.Nakamura and H.Umegaki , On von Neumann measurements in Quantum statistics , Mathematica Japonicae 7 (1960) , 151-157.

[8] J.von Neumann , Die Mathematishe Grandlagen der Quanten Quantenmechanik , Springer , 1932.

[9] 大矢雅則他 , 数理情報科学事典 , 朝倉書店 , 1995年.

[10] C.E.Shannon , A mathematical theory of communication , Bell System Tech. J. 29 (1948) , 398-423 , 628-656.

[11] R.Schatten , Norm Ideal of Completely Continuous Operators , Springer - verlarg , Berlin N.Y. , 1960.

[12] H.Umegaki , On information in operator algebras , Proc. Japan Acad. 37 (1961) , 459-461.

[13] H.Umegaki , Conditonal expectation in an operator algebra , IV (Entropy and

Information) , Kodai Math. Sem. Rep. , 14 (1962) , 59-85.
[14] 梅垣寿春・大矢雅則 , 確率論的エントロピー , 情報理論の函数解析基礎 , 共立出版 1983年 ; 量子論的エントロピー , 同上 共立出版 , 1984年.
[15] 梅垣寿春 , 情報数理の基礎 サイエンス社 , 1994年.